Journal of Cognition and Development

ISSN: (Print) (Online) Journal homepage: https://www.tandfonline.com/loi/hjcd20

Self-regulation in Preschool Children: Factor Structure of Different Measures of Effortful Control and Executive Functions

Sonja Kälin & Claudia M. Roebers

To cite this article: Sonja Kälin & Claudia M. Roebers (2021): Self-regulation in Preschool Children: Factor Structure of Different Measures of Effortful Control and Executive Functions, Journal of Cognition and Development, DOI: 10.1080/15248372.2020.1862120

To link to this article: https://doi.org/10.1080/15248372.2020.1862120

Published online: 03 Jan 2021.

JOURNAL OF COGNITION AND DEVELOPMENT
https://doi.org/10.1080/15248372.2020.1862120

EMPIRICAL ARTICLE

Self-regulation in Preschool Children: Factor Structure of Different Measures of Effortful Control and Executive Functions

Sonja Kälin[a] and Claudia M. Roebers[a]

[a]Department of Psychology, University of Bern, Bern, Switzerland

ABSTRACT

Temperamental effortful control (EC) and executive functions (EF) are two frameworks for studying self-regulation in children. Despite stemming from different research traditions, they show many conceptual and theoretical similarities and their corresponding tasks are often used interchangeably. However, little is known about how and whether the two constructs can be distinguished empirically. The present study aimed to contribute to the investigation of this issue. A sample of 230 preschool children aged 4–6 years were tested with two common behavioral EC tasks and an EC questionnaire. Furthermore, the assessment included common measures of the three subcomponents of EF, namely inhibition, working memory, and shifting. Data were analyzed using correlational and confirmatory factor analyses. In accordance with our hypotheses, we found significant positive correlations between most EC and EF measures, and a single factor model, in which all EC and EF tasks loaded significantly on the underlying factor, was supported by our results. Moreover, this latent construct generalized across gender and age. These findings show that the variety of common EC and EF tasks used in this study all seem to tap similar aspects of self-regulation and therefore support an integrated model of self-regulation encompassing EC and EF. Our results are further considered to be informative for future research using different EC and EF tasks.

Introduction

Imagine a group of preschool children trying to solve a difficult puzzle. Some of the children will persist in solving the task, whereas others will soon give up or be distracted. Should these observed differences in the children's behavior be ascribed to an individual's temperamental regulation, to neurocognitive regulation processes, or maybe to both? The root of this question lies in the fact that in developmental research self-regulation has traditionally been measured with various tasks and questionnaires that are grounded in different research traditions. The temperament-based approach to studying self-regulation in children is known as effortful control (EC). Another approach to studying self-regulation, stemming from the cognitive neuroscience research, is the heterogeneous construct of executive functions (EF). Although a vast amount of research has been conducted regarding EC and EF, only a few studies have systematically compared the two constructs in preschool children and their findings are inconsistent (Blair & Razza, 2007; Neuenschwander, Röthlisberger, Cimeli, & Roebers, 2012; Wolfe & Bell, 2004). Furthermore, there seems to

CONTACT Sonja Kälin sonja.kaelin@psy.unibe.ch Institut für Psychologie, 3012 Bern, Switzerland.

be inconsistency regarding the use of EC and EF tasks in previous research. The question arises how and whether the two constructs EC and EF can be distinguished. Thus, the present study aims to address this question by investigating shared and distinct processes of EC and EF. Given the strong developmental improvements during the preschool years in both EC and EF (Diamond, 2006; Rothbart & Rueda, 2005), we focused on children aged 4–6 years and also looked at possible age-related changes across this age range.

As mentioned above, self-regulation has been approached from different research traditions. EC, on the one hand, is rooted in the research on temperament and is viewed as the self-regulatory component of temperament (Rothbart, Sheese, & Posner, 2007). EC is defined as the efficiency of executive attention involving the ability to inhibit a dominant response to perform a subdominant response, to detect errors, and to plan (Rothbart, Posner, & Kieras, 2006). Factor analyses examining the dimensions of temperament in childhood revealed a broad EC factor including the following subcomponents: attentional focusing, inhibitory control, perceptual sensitivity, persistence, and low intensity pleasure (Rothbart & Bates, 2006). On the other hand, EF are a collection of heterogeneous and top-down neurocognitive processes involved in goal-directed behavior. The common subcomponents of EF are inhibition, working memory, and shifting (Miyake et al., 2000). Inhibition is defined as the ability to inhibit dominant or prepotent responses or behavior. Working memory, also referred to as updating, is defined as short term storage and comprises the ability to actively hold a limited amount of information in mind and mentally manipulating it. Shifting refers to the ability to flexibly shift the focus of attention and being able to adjust to changing task demands or mental sets (Diamond, 2013; Miyake et al., 2000).

Despite stemming from different research traditions, EC and EF have obvious similarities. The two constructs are both identified as top-down processes involved in the regulation of behavior, emotion, and cognition, but are not synonymous with self-regulation (Blair, 2016; Nigg, 2017). Looking at developmental trajectories, in both constructs a rapid increase in performance during preschool years can be observed (Espy, 1997; Kochanska, Murray, Jacques, Koenig, & Vandegeest, 1996), and there is a protracted development into adolescence and adulthood (Rothbart & Rueda, 2005; Zelazo & Müller, 2002). Furthermore, EC and EF share common theoretical features and show overlapping definitions. Inhibition or inhibitory control is viewed as a common component of EC and EF and a central part of both definitions (Miyake et al., 2000; Rothbart & Bates, 2006). Executive attention is assumed to be a common process of both constructs with executive attention being viewed as the attentional component of EF (Blair & Ursache, 2011), and the executive attention network being assumed to underlie EC (Rothbart et al., 2007).

In addition to these theoretical and conceptual similarities, similar – or even the same – tasks are used to assess children's ability to regulate behavior and cognition in both research fields. Certain tasks are even used interchangeably (Zhou, Chen, & Main, 2012). For example, behavioral Go/No-Go tasks, requiring a response to one type of stimuli but not to another, can be found in the EC literature as well as in the EF literature (Diamond & Taylor, 1996; Kochanska & Knaack, 2003). Similarly, Stroop and delay tasks are used in both research traditions and are in some studies labeled as EC and in others as EF tasks (Carlson, 2005; Kochanska, Murray, & Harlan, 2000; Prencipe & Zelazo, 2005). Common EC tasks such as the Whisper task, originally stemming from the temperamental research, are used as an EF measure (e.g., Carlson, Moses, & Breton, 2002; Talwar & Lee, 2008). In fact, EC and EF tasks are sometimes globally labeled as self-regulation tasks or executive control tasks without specifying to which construct they may belong (Denham, Warren-Khot, Bassett, Wyatt, & Perna, 2012; Wiebe, Espy, & Charak, 2008). In conclusion, as EC and EF tasks have originally been developed in two different research areas, it is rather unclear if they measure similar or different aspects of self-regulation. However, if a researcher chooses an EC task to measure EF, it is because they assume this task also measures aspects of EF, and vice versa. Therefore, investigating if commonly used EC and EF tasks capture one or two distinct theoretical constructs is an important contribution to the knowledge on shared processes.

Further similarities concern causes, precursors, and outcomes of EC and EF. EC is defined as constitutionally based, referring to its biological base, and is assumed to be influenced by heredity and maturation. EF are generally not seen as constitutionally based, however, also in EF research a link has been found between genotypic differences and EF performance (Diamond, Briand, Fossella, & Gehlbach, 2004), and investigating twins revealed that EF seem to be highly heritable (Friedman et al., 2008). External factors such as positive family environment and parenting behavior are positively associated with both EC and EF (Eisenberg et al., 2005; Schroeder & Kelley, 2010). EC and EF are both considered predictors for later academic success (e.g., Blair & Razza, 2007) and contribute to positive social development (Eisenberg, Fabes, Guthrie, & Reiser, 2000; Razza & Blair, 2009). At the same time, low performance in either construct has been linked to externalizing problems (Eisenberg et al., 2001; Schoemaker, Mulder, Deković, & Matthys, 2013).

Another similarity between EC and EF concerns the factor structure in preschool children. Using confirmatory factor analysis, it was shown that behavioral EC measures are best described with a single factor (Allan & Lonigan, 2011, 2014). In line with that, another study, including not only EC behavioral measures but also an EC questionnaire, showed that a one-factor solution was most appropriate (Sulik et al., 2010). However, conducting an exploratory factor analysis with various behavioral EC tasks in preschool children resulted in a four-factor solution (Murray & Kochanska, 2002). Regarding EF, different studies using confirmatory factor analyses revealed one unitary executive construct in preschool children (Hughes, Ensor, Wilson, & Graham, 2009; Wiebe et al., 2011; Willoughby, Blair, Wirth, & Greenberg, 2010). There is evidence that the factor structure of EF may depend on task and performance indicator selection (Miller, Giesbrecht, Müller, McInerney, & Kerns, 2012), and the use of CFA might not be the best way of representing the latent construct of EF (Willoughby & Blair, 2016). In conclusion, there are indications that both EC and EF as separate constructs may be best characterized by a single factor structure in preschool years, but choice of tasks and analyses may affect the structure of the resulting constructs.

Despite these major similarities, there is a lack of agreement regarding some of the components of EC and EF. For example, in his recent article, Nigg (2017) outlined the debated role of working memory within temperamental EC. It is unclear whether working memory is part of or closely related to EC (Nigg, 2017; Zhou et al., 2012). In temperamental research, working memory is generally not viewed as a part of EC (Eisenberg, 2017; Rothbart & Bates, 2006). Nevertheless, previous studies found moderate correlations between EC and working memory (Bridgett, Oddi, Laake, Murdock, & Bachmann, 2013), and EC measures and working memory loaded significantly on the same underlying factor (Wiebe et al., 2011). However, these associations do not necessarily prove that EC and working memory belong to the same construct (Eisenberg, 2017). It is possible that the relation between working memory and EC is due to a shared attention component (Nigg, 2017).

Considering the many similarities and overlaps between EC and EF, one would expect to find high correlations between EC and EF measures. However, the few studies investigating EC and EF measures in young children reveal inconsistent results. Whereas some studies including 3–6-year-old children found positive correlations (ranging from small to moderate) between EC and EF behavioral measures (Lin, Liew, & Perez, 2019) as well as between EF tasks and parent- or teacher-rated EC questionnaires (Blair & Razza, 2007; Hongwanishkul, Happaney, Lee, & Zelazo, 2005; Wolfe & Bell, 2004), there is one study finding no significant correlation between the subcomponents of EF and parent-rated EC in 4–8-year-old children (Neuenschwander et al., 2012). Another study with 4-year-olds even showed a moderate negative correlation between working memory and inhibitory control, and a laboratory EC measure (Wolfe & Bell, 2004).

Addressing the question if different measures of EC and EF are indicators of one or more latent factors, to date, there are very few studies including EC measures as well as different subcomponents of EF. Neuenschwander et al. (2012) included all three EF subcomponents and an EC questionnaire, and found two empirically separable and not significantly correlated constructs in kindergarten and 1st grade children. Similarly, in a study with 10–15-year-old children that included different questionnaires tapping EC and behavioral EF inhibition measures, a single factor model showed a moderate fit, but none of the performance-based measures significantly loaded on the underlying factor (Samyn, Roeyers, Bijttebier, Rosseel, & Wiersema, 2015). Thus, the idea that EC questionnaires and EF tasks measure the same underlying construct could not be supported. However, a very recent study of Lin et al. (2019) with 4–6-year-old children included performance-based measures of EF inhibition and EC and using confirmatory factor analysis, a one-factor model of self-regulation was supported. This large overlap between EC and EF has also been supported in an adult sample (Bridgett et al., 2013). In summary, despite the theoretical and conceptual similarities between EC and EF, the empirical evidence is less clear with mostly small to moderate associations between different measures of EC and EF. Not much is known about the factor structure of EC and EF measures and to our best knowledge, there is no study looking at the factor structure of a broad assessment including common EF measures of the three subcomponents (inhibition, working memory and shifting) as well as EC assessed with both behavioral tasks and a questionnaire.

Another interesting aspect when studying self-regulation is the finding of gender differences in various self-regulation measures. Previous research showed differences between boys and girls in parent-report measures of self-regulation favoring girls (Murphy, Eisenberg, Fabes, Shepard, & Guthrie, 1999; Raffaelli, Crockett, & Shen, 2005). Likewise, a meta-analysis investigating gender differences in temperament found that EC showed large gender differences favoring girls when relying on parent and teacher reports (Else-Quest, Hyde, Goldsmith, & Van Hulle, 2006). Gender differences were also found in behavioral measures of EC and EF (e.g., Carlson & Moses, 2001; Kochanska et al., 2000). Nevertheless, various studies assessing EC and EF using behavioral tasks or questionnaires do not find any differences between boys and girls (e.g., Carlson et al., 2002; Davis, Bruce, & Gunnar, 2002). Similar inconsistencies can be found when looking at gender differences in factor analyses. Investigating measurement invariances, some studies showed that the constructs of EC and EF appear to be similar across gender (Allan & Lonigan, 2011, 2014; Wiebe et al., 2011), whereas in other studies the factor loadings and intercepts differed for boys and girls with higher intercepts in girls for behavioral EC measures and for teacher-rated EC (Denham et al., 2012; Sulik et al., 2010). Thus, previous findings regarding gender differences in EC and EF measures failed to show a consistent pattern, motivating the exploration of gender differences in the present study.

In summary, against the background of the above reviewed findings, a more thorough investigation of the EC and EF constructs and their measurements seems warranted. Recently, there has been an increasing pledge for a more integrative view on self-regulation (Nigg, 2017; Zhou et al., 2012) and the idea was put forward to perceive EC and EF as complementary rather than incompatible (Liew, 2012). Zhou et al. (2012) concluded that it is difficult to distinguish EC and EF as global constructs, although it is possible to identify distinctions among the components, processes and measures of EC and EF. In order to reduce overlaps and confusion in future research using EC and EF measures, the possibility of an integrated theory of self-regulation encompassing EC and EF should be taken into account and further investigated. Using confirmatory factor analyses, the present study aimed to contribute empirical data to this issue by examining the interrelations between different EC and EF tasks and whether the different assessments can be linked to one or more underlying factors. In view of the current knowledge, we hypothesized to find positive correlations between different measures of EC and EF. In line with a possible integrated model, we further expected variances in different EC and EF measures to be best described by a single factor model. Due to expected age effects between 4–6-year-olds, we investigated age related changes and additionally, the role of gender was examined to find out more about the relationship and structure of different self-regulation assessments in boys and girls.

Methods

Participants

The final study sample consisted of $N = 230$ children (50% female) between 4 and 6 years of age ($M = 68.5$ months, $SD = 7.4$ months) attending kindergarten. Children were recruited from public kindergartens in different urban and rural regions of the German speaking part of Switzerland. They were predominately Caucasian from lower and upper middle-class families, reflecting the characteristics of the local community. All included participants were fluent in the German language, 27 children had to be excluded due to insufficient German language skills. Technical problems led to the exclusion of 13 more children, 5 children were excluded because they were older than 83 months and 8 children noticed the hidden camera in the Puzzle Box task and were therefore excluded from all analyses. There was no significant difference between children included in the study sample and the excluded children regarding age and performance in EC and EF tasks. Participation was voluntary and parents gave their written informed consent. Children themselves agreed to participate before the testing. The study was approved by the faculty's ethics committee (Approval No. 2017-04-00006).

Procedure

Children were individually assessed by trained experimenters in a quiet room at the children's kindergarten. Each child was tested in three sessions each lasting about

30 minutes. The three sessions were counterbalanced between children and gender, with each session including three tasks in a fixed order. After completion of all sessions, children were thanked for their cooperation and received a small gift.

Measures

Effortful control

Two often used behavioral EC tasks and a common temperament questionnaire were chosen to assess diverse aspects of the EC construct. One of the measures was the Puzzle Box, which is assumed to assess persistence and inhibitory control (Spinrad, Eisenberg, & Gaertner, 2007). Children were instructed to try to assemble a wooden puzzle without looking at it. The puzzle was placed in a wooden box behind a cloth that covered the front and children had to slip their arms through sleeves to access the puzzle. The cloth could be lifted up so that the child could easily cheat by looking at the puzzle. Children were told that they have 5 minutes to finish the puzzle and if they completed it earlier they should call the experimenter. The experimenter left the room and children's behavior was recorded using hidden cameras. Seconds during which the child displayed persistence (e.g. trying to solve the puzzle), cheating (e.g. lifting the cloth), and off-task (e.g. walking around the room) behavior were double-coded; interrater reliability was .99. The dependent variable was the accumulated time of persistence divided by the total task-time. We focused on the persistence component because persistence is by definition part of the EC construct.

The Whisper task is part of Kochanska's battery for assessing EC and is considered to measure inhibitory control (Kochanska et al., 1996). Children were asked to whisper the names of well-known characters of cartoons or children's books. After making sure that children knew how to whisper by whispering their own names, 12 pictures were presented consecutively to them. Each trial was coded as 0 (shouts), 1 (speaks normal or changes from normal to whisper), or 2 (whispers). Behavior during the task was video recorded and all answers were double-coded; the interrater reliability was .99.

Since EC is traditionally measured with questionnaires, we added a parent-rated questionnaire using the very short form of the Children's Behavior Questionnaire (CBQ; German translation; Putnam & Rothbart, 2006). The very short form of the CBQ consists of 36 items; 12 of the items compose the EC scale (3 items per subscale: inhibitory control, attention focusing, low-intensity pleasure, and perceptual sensitivity). Parents were asked to rate the behavior of their child on a 7-point Likert scale ranging from *extremely true* to *extremely untrue*. A *not applicable* response option was also present and to be used in case the child had not been observed in the situation described. The EC score was derived by calculating the mean of the parents' rating of all 12 items. Higher scores reflect better EC. Internal consistency of our sample was .66, comparable with the internal consistency (ranging from .62 –.78) obtained in three different samples when developing the very short form of the CBQ (Putnam & Rothbart, 2006).

Executive functions

Although research does not clearly show separable components of EF in young children (e.g., Lee et al., 2012), the three common subcomponents inhibition, working memory, and shifting were assessed in the current study in order to achieve a broad and comprehensive measurement. Thereby, each EF task was assumed to assess mainly one of the three EF components.

Inhibition was assessed using an adapted and computerized version of the Fruit Stroop task (Archibald & Kerns, 1999) consisting of three blocks. Each block included 24 trials presented on a tablet screen, with each trial consisting of a target stimulus appearing for one second on the first screen and a response screen with four different colors (red, green, blue, yellow). Children had to choose one of the four colors as quickly as possible by touching the screen. In the first block, colored squares were displayed as target stimuli (baseline condition) and in the second block, four different fruits or vegetables in their original color were presented as target stimuli (congruent condition). Children had to choose the same color on the response screen. In the third block, the same fruits and vegetables were displayed in incongruent colors and children were asked to choose the original color on the response screen (e.g. a blue strawberry was presented and children had to press the red color; incongruent condition). The dependent variable was a combined score including the accuracy and the reaction time (reverse scored) of the incongruent condition (z-value of accuracy + reversed z-value of mean reaction time).

The backwards Color Span (Zoelch, Seitz, & Schumann-Hengsteler, 2005) was used as a measurement of working memory. Children were told a cover story about a dwarf who loses colored discs. Sequences of colored discs were then presented on a tablet screen and children were asked to recall the colors in reverse order. Each color appeared on the screen for one second. Only colors with monosyllabic names (in German) were selected. After three practice trials, each child started with two-item sequences. Sequence length increased by one item if the child correctly recalled three of the six sequences on a particular level. The dependent variable was the total number of correctly recalled sequences.

The EF subcomponent shifting was measured with a modified Dimensional Change Card Sorting task (DCCST; Carlson, 2005; Zelazo, Müller, Frye, & Marcovitch, 2003). Children were introduced to three boxes with slots cut on the top and each box displayed a target card with a certain colored shape. The task consisted of three conditions, each including one practice trial. After making sure that the child was familiar with all the colors and shapes, the children were asked to sort six cards according to color as quickly as possible (condition 1). In the second condition, children were asked to sort six cards according to shape as quickly as possible. In the third condition, children were introduced to cards with stars. Children were given 18 cards (5 star, 13 non-star) and told to sort all cards with a star according to shape and all cards without a star according to color as fast as possible. The dependent variable was a combined score including the total number of errors and the total time in seconds of the third condition ((errors + 1) * time). The variable was reversed scored, with a higher value representing better performance.

Statistical analyses

Data were analyzed using the software Jamovi 1.1 (The jamovi project, 2019) which is running on R (R Core Team, 2018). In the Stroop task, outliers regarding the reaction times of each subject were excluded if they deviated +- 3 standard deviations from the subject's mean reaction time. This applied to 2% of all reaction times. Additionally, in all dependent variables, scores that were higher or lower than 3 standard deviations from the sample's mean were set to 3 standard deviations. This concerned 0.4% of all values in the data set. No multivariate outliers were found, as assessed by the Mahalanobis distance ($p > .001$).

The scores of the Whisper task were negatively skewed. Transforming the values using a reflected logarithm to normalize the skewed distribution yielded the same results and in order to ease interpretation of results, we will report analyses of the raw scores exclusively. Children with one or more missing variables were excluded from all analyses (N = 48), which led to a final sample of 230 children. It is important to note that 27 of the excluded children were not able to properly understand the task instructions due to insufficient German language skills, which led to data that were not valid enough to use data imputation.

Confirmatory factor analyses were conducted using IBM SPSS Amos 25 software. Model fit to the data was assessed using multiple fit measures, namely the χ^2 value, the comparative fit index (CFI), the root mean square error of approximation (RMSEA) and its 90% confidence interval (90% CI), and the standardized root mean square residual (SRMR). CFI indexes higher than .95, RMSEA values less than .06, and SRMR values below .05 indicate a good fit (Hu & Bentler, 1999). To test for gender differences, multi-group analyses were conducted to evaluate different levels of measurement invariance. Measurement invariance was tested using the χ^2 difference test and the cutoff criteria suggested by Chen (2007): a change of >-.005 in CFI, supplemented by a change of >.010 in RMSEA indicates violation from invariance.

Results

Performance in EC and EF measures

Descriptive data for all task variables of the current study are presented in Table 1. The mean performance in all assessments separated by gender is shown in Table 2. ANOVAs were run to determine gender differences in all EC and EF measures. There was a significant difference between girls` and boys` performance for inhibition, $F(1, 228) = 9.05$, $p = .003$, with this difference yielding a small effect size ($\eta^2 = 0.038$). More detailed analyses revealed that this difference was due to a significantly better performance of girls regarding the accuracy in the incongruent condition of the Stroop task, whereas no significant gender difference was found for reaction time. Furthermore, there was

Table 1. Descriptive statistics for all task variables.

		Mean	SD	Min	Max.
Executive Functions					
	Stroop	.03	1.36	−3.55	2.64
	Color Span	7.17	4.06	0	17
	DCCST	362	236	47	1090
Effortful Control					
	Whisper	20.2	4.69	7	24
	Puzzle Box	0.56	0.29	0.0	1.0
	CBQ	5.34	.68	3.30	6.83

N = 230. Reaction times of the Stroop Task are reverse scored and combined with accuracy. DCCST = Dimensional change card sorting task. CBQ = Children's behavior questionnaire.

Table 2. Comparison of performance in all self-regulation tasks between boys and girls.

	Mean		SD		F	*p*	η^2
	m	f	m	f			
Executive Functions							
Stroop	−.23	.30	1.35	1.33	9.05	.003	.038
Color Span	7.30	7.03	4.25	3.88	.25	.616	.001
DCCST	368	356	254	218	.15	.695	.001
Effortful Control							
Whisper	20.1	20.3	4.77	4.62	.11	.736	.000
Puzzle Box	0.55	0.57	0.29	0.30	.43	.511	.002
CBQ	5.24	5.45	0.68	0.67	5.45	.020	.023

N = 230. Male (m) = 115; Female (f) = 115. DCCST = Dimensional change card sorting task. CBQ = Children's behavior questionnaire. Reaction times of the Stroop Task are reverse scored and combined with accuracy.

a significant difference between girls and boys in the EC questionnaire, $F(1, 228) = 5.45$, $p = .02$, again yielding a small effect size ($\eta^2 = 0.023$). All other EC and EF measures showed no significant difference between girls and boys.

Correlational analyses

To investigate the relationship between different EC and EF measures, Pearson correlations were computed. The intercorrelations among all dependent variables, the correlation with age as well as the partial correlations after controlling for age are shown in Table 3. There were significant correlations among all EF tasks. However, after controlling for age the correlation between the Stroop and the Color Span was no longer significant. There was a significant correlation between the two EC behavioral tasks (Whisper, Puzzle Box) as well as between the EC questionnaire and the Puzzle Box. Regarding the intercorrelations between EC and EF measures, performance in the Whisper task and persistence in the Puzzle Box task were significantly correlated with the Stroop task (inhibition). Additionally, the Puzzle Box task and the EC subscale of the CBQ were significantly correlated with the EF shifting task (DCCST). All EF and EC measures correlated significantly with age except for the EC questionnaire. Looking at the correlations separately for girls and boys and comparing them using Fisher's Z showed no significant differences in either of the EC and EF variables.

Table 3. Correlations among executive functions and effortful control variables, and age.

	1	2	3	4	5	6	Age
Executive Functions							
1. Stroop	-	.07	.21**	.25***	.18**	.11	.45***
2. Color Span	.14*	-	.22**	.10	.08	.07	.17**
3. DCCST	.31***	.26***	-	.11	.24***	.17**	.28***
Effortful Control							
4. Whisper	.33***	.14*	.17*	-	.18**	.09	.25***
5. Puzzle Box	.24***	.11	.28***	.21**	-	.14*	.18**
6. CBQ	.08	.07	.15*	.08	.13*	-	−.04

N = 230. DCCST = Dimensional change card sorting task. CBQ = Children's behavior questionnaire. Reaction times of the Stroop and DCCST are reverse scored. The upper right triangle represents the partial correlation coefficients after controlling for age.

* *p* <.05, ** *p* <.01, *** *p* <.001.

Confirmatory factor analyses

To investigate whether the different EC and EF measures can be linked to one or more latent variables, two hypothesized models were calculated. Model 1 was a single-factor model in which all self-regulation measures loaded on a single factor. Model 2 was a two-factor model in which the first factor consisted of all EF tasks, and the second factor consisted of all EC tasks and the EC questionnaire. The single factor model provided a good fit to the data with CFI = .98, RMSEA = .030 with 90% CI [.00, .084], and SRMR = .037. The two-factor model as well provided a good fit to the data with CFI = .97, RMSEA = .039 with 90% CI [.00, .092], and SRMR = .036. In both models, all variables loaded significantly on the underlying factor (see Figure 1). However, in both factor models, the EC questionnaire showed only a small factor loading that was below the .30 loading cutoff criterion (Bowers et al., 2010). The between-factor correlation in the two-factor model was .95 and highly significant. Comparing the model fit of the two models using the χ^2 difference test resulted in χ^2 (1, N = 230) = 0.12, p = .73, revealing no significant difference between the one- and two-factor model. Additionally, given the very high correlation (r = .95) between the two latent

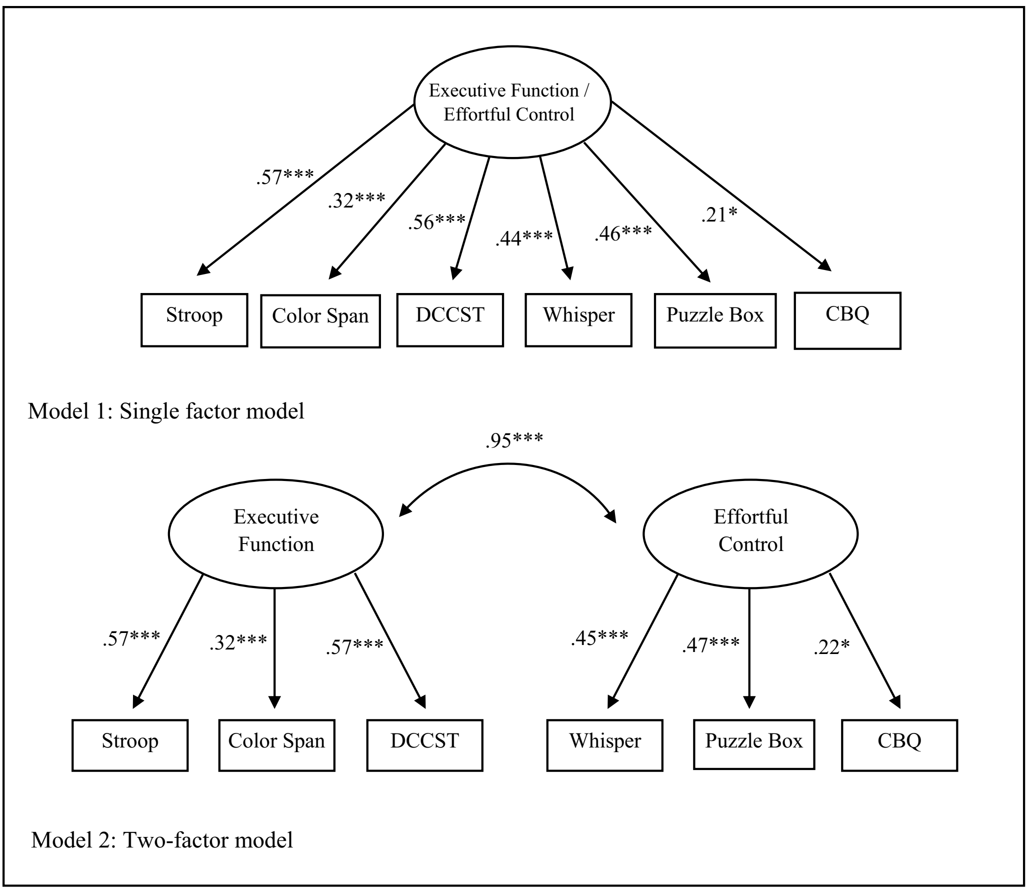

Figure 1. The two hypothesized factor models. Standardized factor loadings are shown. DCCST = Dimensional change card sorting task. CBQ = Children's behavior questionnaire.
* p <.05, ** p <.01, *** p <.001.

variables in the two-factor model, it can be questioned if the two factors had much unique explanatory power. Due to this high correlation and by parsimony, the one-factor model was selected as the better-fitting model.

After controlling for age (see Figure 2), both hypothesized factor models still revealed a good model fit with CFI = .99, RMSEA = .017 with 90% CI [.00, .078], and SRMR = .029

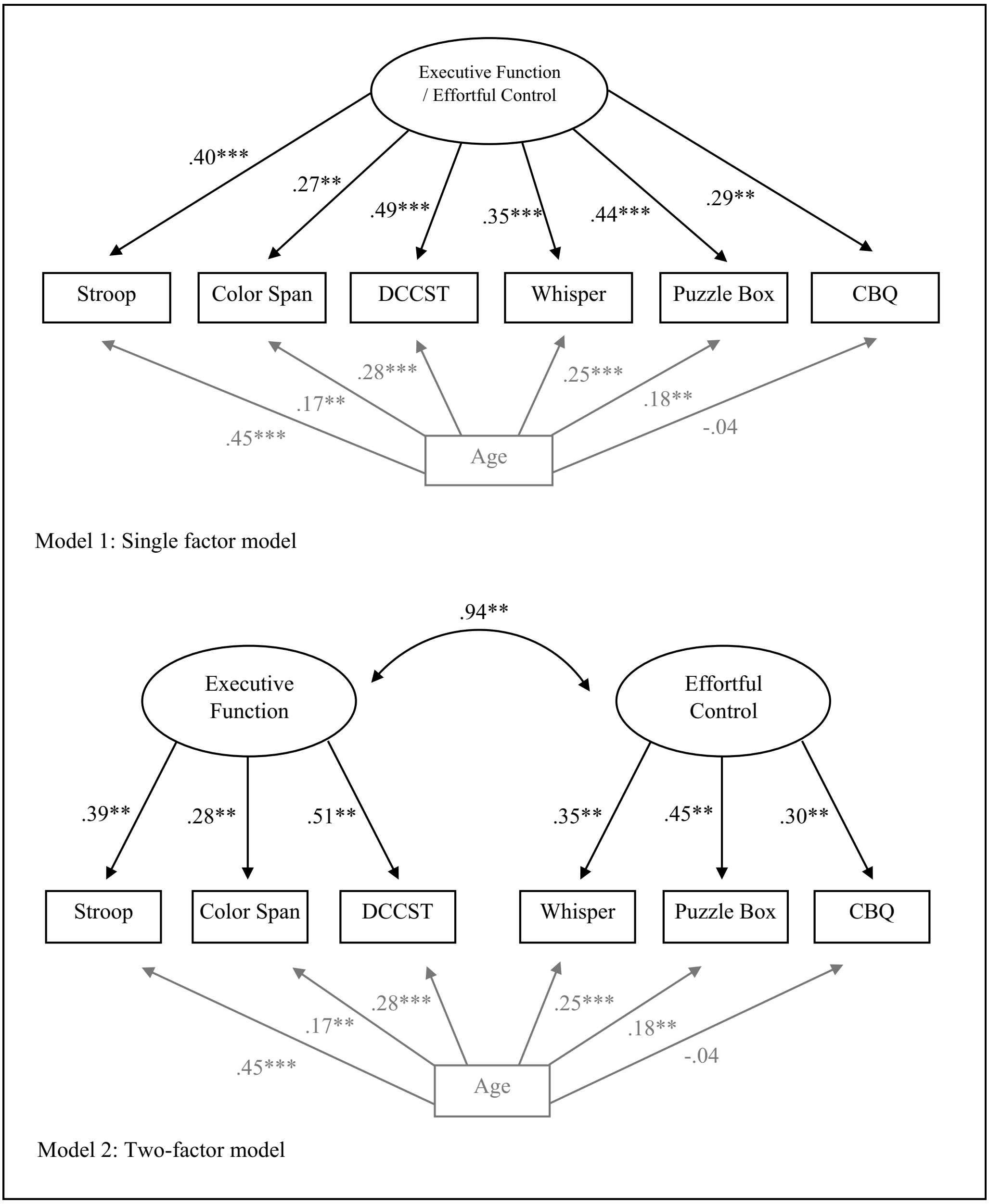

Figure 2. The two hypothesized factor models with age (in months) included as control variable. Standardized factor loadings are shown. DCCST = Dimensional change card sorting task. CBQ = Children's behavior questionnaire. * p <.05, ** p <.01, *** p <.001.

for the one factor model and CFI = .99, RMSEA = .029 with 90% CI [.00, .086], and SRMR = .029 for the two-factor model. All variables still showed a significant factor loading. In addition, we built two age groups (5- and 6-year-olds, N = 112 and 89) and ran both hypothesized factor models for each age group separately to investigate if there were age-related changes in the factor structure or if the better-fitting model was equivalent for both age groups. In both age groups, comparing the model fit of the two models using the χ^2 difference test revealed no significant difference between the one- and two factor model (5-year-olds: χ^2 (1, N = 112) = 0.01, p = .92; 6-year-olds: χ^2 (1, N = 89) = 0.12, p = .73). Moreover, the correlation between the two latent variables was very high in the group of 5-year-olds (r = 1.00) as well as in the group of 6-year-olds (r = .91). Consequently, in both age groups the one-factor model was chosen as better-fitting model.

Due to measurement differences between the behavioral measures and the questionnaire, we additionally calculated a one- and two-factor model without the EC questionnaire. There was no significant difference between these two models when comparing the model fit using the χ^2 difference test (χ^2 (1, N = 230) = 0.05, p = .83), and the correlation between the two latent variables in the two-factor model remained about equally high (r = .97) when compared with the model including the parental paper-and-pencil questionnaire for EC.

Gender and age invariance

Multi-group analyses were conducted to find out if the model fit of our single factor model depended on gender or age (Table 4). In a first step, metric invariance (whether the magnitude of factor loadings is the same) was tested. Regarding gender, setting equality constraints on factor loadings did not result in a significantly worse model fit based on the χ^2 difference test and the difference in CFI and RMSEA. Testing scalar invariance (whether the intercept of the regression relating each item to its factor is the same), the χ^2 difference test showed a significant result and also the CFI and RMSEA differences indicated that invariance of the intercepts was not achieved. This suggests that at least one item intercept differed significantly between girls and boys. Further analyses revealed that the lack of scalar invariance applied to EF inhibition measured with the Stroop task and the EC subscale of the CBQ. For all other tasks, scalar invariance was attained. Testing metric invariance across the two age groups of 5- and 6-year-olds showed a worse model fit after constraining the factor loadings to be equal, which was indicated by a significant χ^2 difference test and significant CFI and RMSEA differences. More detailed analyses revealed different factor

Table 4. Summary of fit statistics for testing gender and age invariance of the single factor model.

	χ^2	df	CFI	RMSEA	$\Delta\chi^2$	df	p	ΔCFI	ΔRMSEA
Boys vs. girls (N = 230, 115 vs. 115)									
Unconstrained	17.55	18	1.00	.000	-	-	-	-	-
Factor loadings invariant	19.19	24	1.00	.000	1.64	6	.95	.000	.000
Intercepts invariant	33.78	30	.955	.024	14.59	6	.02	−.045	−.024
5-year-olds vs. 6-year-olds (N = 201, 112 vs. 89)									
Unconstrained	10.32	18	1.00	.000	-	-	-	-	-
Factor loadings invariant	25.50	24	.975	.018	15.18	6	.02	−.025	−.018
Intercepts invariant	56.22	30	.559	.066	30.72	6	.00	−.416	−.048

CFI = Comparative fit index; RMSEA = Root mean square error of approximation.

loadings of the two EF measures Stroop and DCCST. Furthermore, constraining the intercepts to test scalar invariance also resulted in a significant χ^2 difference test and significant CFI and RMSEA differences, which indicated a worse model fit. Further analyses revealed that scalar invariance was not attained for any EC and EF measures except the CBQ.

Discussion

The main goal of this study was to investigate if EC and EF can conceptually be viewed as one construct and to what extent different EC and EF tasks share common processes. Furthermore, we aimed to investigate gender and age differences. In view of the question of an integrated model of self-regulation encompassing EC and EF, we hypothesized a) to find positive associations between different EC and EF measures and b) that EC and EF were better viewed as a global unidimensional construct. Overall, our findings showed that this was the case.

The results of the present study indicate that different EC and EF measures were best conceptualized as a unitary construct and all variables loaded significantly on the underlying factor. Consequently, the EC and EF measures used in this study seem to share substantial amounts of variance. Although one could expect the questionnaire not to load on the same factor as behavioral EC and EF measures (Samyn et al., 2015), our results showed that also the questionnaire loaded significantly on the underlying factor. However, the factor loading was lower than for most of the behavioral tasks. Keeping in mind that measurement differences may appear between behavioral assessments and parental reports, and possibly account for the lower factor loading, the significant factor loading still suggests that the questionnaire can be considered to measure the same underlying construct. An explanation for the diverging results between our approach and the study of Samyn et al. (2015) could lie in the older age group and the use of different EC questionnaires. It is possible that in our group of preschool children the EC questionnaire and behavioral EF measures tap the same underlying construct, whereas later in development the two types of measures tap different kinds of information.

Our finding of a single factor model is in contrast to Neuenschwander et al. (2012) finding two separate factors for EC and EF. However, our results extend this previous work by examining EC not only with a questionnaire but also with behavioral measures. Another difference is that Neuenschwander et al. (2012) included the EC questionnaire in the model by using three parcels of items. It cannot be excluded that we would have obtained different results if we had used the same method. Nevertheless, the inclusion of different EC measures allows to get a broader conception of the structure of different self-regulation tasks. Similar to Lin et al. (2019), our study involved multiple measures of EC and EF, and our results indicated that the two constructs cannot be clearly distinguished. In addition to the study of Lin et al. (2019), we included an EC questionnaire and EF was assessed not only with inhibition tasks but also a working memory and a shifting task. Our findings demonstrate that various EC and EF tasks are not pure measures of their theoretical construct, but share variance with the other, related construct. Firstly, this provides support for an integrated model of self-regulation encompassing EC and EF, as it was proposed by Zhou et al. (2012). Secondly, despite the fact that our tasks traditionally stem from different research traditions, they all seem to measure similar aspects of self-regulation.

The unidimensional EC and EF construct in our study was established in a sample of preschool children. It is important to note that in preschool years both EC and EF undergo dramatic changes (Diamond, 2006; Rothbart & Rueda, 2005), which is supported by the age-related improvements in performance that we found in most of our EF and EC measures. Only in the CBQ, children did not improve with age, which can easily be explained by the fact that parents usually take age into account when filling out a questionnaire. Investigating the factor structure of our EC and EF measures separately for 5- and 6-year-olds revealed that in both age groups, the one factor model was the better-fitting model. Similarly, using a longitudinal design and including only EF tasks, Hughes et al. (2009) found that a single latent EF factor provided a good fit to the data for children at both 4 and 6 years of age. Interestingly, investigating the factor structure of EF in older children (8–13 years) revealed three interrelated factors, similar to the factor structure obtained in an adult sample (Lehto, Juujärvi, Kooistra, & Pulkkinen, 2003; Miyake et al., 2000). These studies indicate that EF may be characterized by a single factor structure in preschool years and become increasingly differentiated with age. It is possible that our latent construct consisting of EF and EC also differentiates with age and progressive development.

According to our hypothesis, we found significant positive links between most EC and EF measures. However, neither of the EC measures was significantly associated with working memory. These findings are especially interesting regarding the debated role of working memory within temperament. Working memory is generally not viewed as a part of EC which seems to be supported by our correlational results. Nevertheless, the Color Span task and all three EC measures loaded significantly on the same underlying factor, as was found before by Wiebe et al. (2011). This indicates that working memory and EC seem to tap similar facets of self-regulation, with a common attentional component being a candidate process (Nigg, 2017). Future research might perform in-depth analyses to further investigate a potential involvement of working memory in both EC and EF tasks.

Another difference between EC and EF should be highlighted in view of our findings. Although both EC and EF are assumed to be influenced by the interaction of genes and environment, temperament traits tend to show consistency across situations and stability over time (Rothbart & Rueda, 2005), which has also been supported by studies with preschool children (e.g., Kochanska et al., 2000). EF has been proven to be relatively responsive to changes by interventions or trainings (Diamond & Lee, 2011). An early intervention regarding EF can be crucial to change a child's developmental trajectory and can help to reduce the gap between children with better and worse EF (Diamond, 2013). To date, not much is known about the responsiveness of EC to training or interventions, as there are only few such investigations. Interventions in the field of temperamental EC have mainly focused on indirect effects of trainings, such as attentional training with the idea of attention being a neural substrate of EC (Posner, Rothbart, & Rueda, 2008) or by fostering parental skills in order to facilitate children's EC (Chang, Shaw, Dishion, Gardner, & Wilson, 2015). Considering our findings of a unidimensional construct comprising EC and EF, it would be interesting to conduct training and intervention studies investigating different aspects of self-regulation by including various EC and EF measures.

A further distinction between EC and EF has sometimes been made by means of emotional or neutral contexts with EC traditionally focusing more on emotional contexts and EF focusing on neutral contexts. This distinction can also be labeled as "hot" and "cool" which goes back to Metcalfe and Mischel (1999). Hot tasks are defined as tasks with a salient emotional component and cool tasks as tasks without an emotional component, but with more cognitive and abstract problem-solving (Zelazo & Müller, 2002). In our study, the two behavioral EC tasks could generally be labeled as hot tasks involving more emotional components whereas the EF tasks would rather count as cool tasks without any circumscribed emotional aspects involved. Our finding that EC and EF were best represented as a unidimensional construct also seems interesting in light of this distinction between hot and cool components. It suggests that even measures with differing emotional salience have a lot in common, which is also supported by previous research looking at the factor structure of hot and cool aspects of self-regulation tasks (Allan & Lonigan, 2011, 2014; Denham et al., 2012; but see Willoughby, Kupersmidt, Voegler-Lee, & Bryant, 2011, for a two factor model). Interestingly, to date, the cool component of self-regulation has often only been measured with different inhibition tasks, while neglecting the EF subcomponents working memory and shifting. Therefore, our findings extend these previous findings by adding working memory and shifting to the cool component of self-regulation and still finding a unitary construct.

Regarding gender differences in our factor model, full metric invariance and partial scalar invariance was established, meaning that our latent construct generalized across gender. As is supported by the literature investigating the factor structure of EC or EF separately (Allan & Lonigan, 2014; Sulik et al., 2010; Wiebe et al., 2011), our finding of metric invariance indicates that the factor loadings of our single model encompassing EC and EF were not significantly different between boys and girls. Therefore, each item contributed to our latent construct in a similar way across gender. However, our results only supported partial scalar invariance. More specifically, the intercept for the EF inhibition task was larger in girls than in boys, meaning that mean differences in our latent EC and EF construct do not capture all mean differences in the shared variance of the Stroop task (Putnick & Bornstein, 2016). Others also found differences in intercepts between males and females in an inhibitory control task (Sulik et al., 2010). As scalar invariance across gender regarding the Stroop task and the CBQ could not be established in our sample, the comparison of mean levels between girls and boys is not justified and cannot be interpreted. The comparison of the means of the remaining EC and EF measures is valid and revealed no significant differences between boys and girls. To sum up, although more pronounced gender differences could have been expected considering previous EC and EF research (Carlson & Moses, 2001; Else-Quest et al., 2006), overall, our results do not support substantial group differences between boys and girls.

There were limitations in this study that should be considered when interpreting the findings. Firstly, a methodological issue is that we included behavioral measures and a questionnaire to assess EC, but did not include a questionnaire to assess EF. Thus, we cannot rule out that our findings are influenced by measurement differences. However, as EC has traditionally often been assessed with questionnaires, this multi-method approach also entails strengths. Additionally, as reported, we tested our hypothesis with a one- and two-factor model without the EC questionnaire and this led to the same main findings. Secondly, we were able to establish convergent validity in our study, but could not demonstrate discriminant validity. Therefore, we cannot provide evidence to what extent our EC and EF measures share variance with other constructs. Further research might consider assessing additional tasks to investigate discriminant validity. Thirdly, it is important to keep in mind that our findings may depend on the specific measures we used for each construct and cannot unconditionally be extended to all other existing EC and EF measures. Furthermore, our measures only showed weak to moderate correlations with each other.

In conclusion, our results indicate that a variety of common EC and EF tasks load onto a single latent construct and that this construct behaves in a similar way across gender. The present findings seem relevant not only from a conceptual perspective, but also regarding the use of different self-regulation tasks. On a conceptual basis, our results support the idea of an integrated model of self-regulation encompassing EC and EF, and could be beneficial to clarify the use of concepts and terminologies in future research. Regarding the question of which EC and EF task is used in which context in the literature, our results demonstrate that the different EC and EF measures included in our study share variance and can be assumed to measure similar aspects of self-regulation. Our results are promising and future work should consider examining other behavioral EC and EF measures as well as including different EC and EF questionnaires. Furthermore, it would be interesting to investigate the possible role of an attention component linking EC and EF by including such a measurement.

Acknowledgments

We would like to thank Nike Tsalas for her valuable contributions in setting up the study. We further thank all participating children and their parents as well as the kindergarten teachers for their cooperation. We also gratefully acknowledge the help of our research assistants and master students concerning data collection and behavioral coding.

Disclosure statement

No potential conflict of interest was reported by the authors.

References

Allan, N. P., & Lonigan, C. J. (2011). Examining the dimensionality of effortful control in preschool children and its relation to academic and socioemotional indicators. *Developmental Psychology*, *47*(4), 905–915. doi:10.1037/a0023748

Allan, N. P., & Lonigan, C. J. (2014). Exploring dimensionality of effortful control using hot and cool tasks in a sample of preschool children. *Journal of Experimental Child Psychology*, *122*, 33–47. doi:10.1016/j.jecp.2013.11.013

Archibald, S. J., & Kerns, K. A. (1999). Identification and description of new tests of executive functioning in children. *Child Neuropsychology*, *5*(2), 115–129. doi:10.1076/chin.5.2.115.3167

Blair, C. (2016). The development of executive functions and self-regulation. In R. F. Baumeister & K. D. Vohs (Eds.), *Handbook of self-regulation: Research, theory, and applications* (3rd ed., pp. 417–439). New York, NY: Guilford Press.

Blair, C., & Razza, R. P. (2007). Relating effortful control, executive function, and false belief understanding to emerging math and literacy ability in kindergarten. *Child Development*, *78*(2), 647–663. doi:10.1111/j.1467-8624.2007.01019.x

Blair, C., & Ursache, A. (2011). A bidirectional model of executive functions and self-regulation. In R. F. Baumeister & K. D. Vohs (Eds.), *Handbook of self-regulation: Research, theory, and applications* (2nd ed., pp. 300–320). New York, NY: Guilford Press.

Bowers, E. P., Li, Y., Kiely, M. K., Brittian, A., Lerner, J. V., & Lerner, R. M. (2010). The five Cs model of positive youth development: A longitudinal analysis of confirmatory factor structure and measurement invariance. *Journal of Youth and Adolescence*, *39*(7), 720–735. doi:10.1007/s10964-010-9530-9

Bridgett, D. J., Oddi, K. B., Laake, L. M., Murdock, K. W., & Bachmann, M. N. (2013). Integrating and differentiating aspects of self-regulation: Effortful control, executive functioning, and links to negative affectivity. *Emotion*, *13*(1), 47–63. doi:10.1037/a0029536

Carlson, S. M. (2005). Developmentally sensitive measures of executive function in preschool children. *Developmental Neuropsychology*, *28*(2), 595–616. doi:10.1207/s15326942dn2802_3

Carlson, S. M., & Moses, L. J. (2001). Individual differences in inhibitory control and children's theory of mind. *Child Development*, *72*(4), 1032–1053. doi:10.1111/1467-8624.00333

Carlson, S. M., Moses, L. J., & Breton, C. (2002). How specific is the relation between executive function and theory of mind? Contributions of inhibitory control and working memory. *Infant and Child Development: An International Journal of Research and Practice*, *11*(2), 73–92. doi:10.1002/icd.298

Chang, H., Shaw, D. S., Dishion, T. J., Gardner, F., & Wilson, M. N. (2015). Proactive parenting and children's effortful control: Mediating role of language and indirect intervention effects. *Social Development*, *24*(1), 206–223. doi:10.1111/sode.12069

Chen, F. F. (2007). Sensitivity of goodness of fit indexes to lack of measurement invariance. *Structural Equation Modeling*, *14*(3), 464–504. doi:10.1080/10705510701301834

Davis, E. P., Bruce, J., & Gunnar, M. R. (2002). The anterior attention network: Associations with temperament and neuroendocrine activity in 6-year-old children. *Developmental Psychobiology: The Journal of the International Society for Developmental Psychobiology*, *40*(1), 43–56. doi:10.1002/dev.10012

Denham, S. A., Warren-Khot, H. K., Bassett, H. H., Wyatt, T., & Perna, A. (2012). Factor structure of self-regulation in preschoolers: Testing models of a field-based assessment for predicting early school readiness. *Journal of Experimental Child Psychology*, *111*(3), 386–404. doi:10.1016/j.jecp.2011.10.002

Diamond, A. (2006). The early development of executive functions. *Lifespan Cognition: Mechanisms of Change*, *210*, 70–95.

Diamond, A. (2013). Executive functions. *Annual Review of Psychology*, *64*(1), 135–168. doi:10.1146/annurev-psych-113011-143750

Diamond, A., Briand, L., Fossella, J., & Gehlbach, L. (2004). Genetic and neurochemical modulation of prefrontal cognitive functions in children. *American Journal of Psychiatry*, *161*(1), 125–132. doi:10.1176/appi.ajp.161.1.125

Diamond, A., & Lee, K. (2011). Interventions shown to aid executive function development in children 4 to 12 years old. *Science*, *333*(6045), 959–964. doi:10.1126/science.1204529

Diamond, A., & Taylor, C. (1996). Development of an aspect of executive control: Development of the abilities to remember what I said and to "Do as I say, not as I do". *Developmental Psychobiology*, *29*(4), 315–334. doi:10.1002/(SICI)1098-2302(199605)29:4<315::AID-DEV2>3.0.CO;2-T

Eisenberg, N. (2017). Commentary: What's in a word (or words)–on the relations among self-regulation, self-control, executive functioning, effortful control, cognitive control, impulsivity, risk-taking, and inhibition for developmental psychopathology–reflections on Nigg (2017). *Journal of Child Psychology and Psychiatry*, *58*(4), 384–386. doi:10.1111/jcpp.12707

Eisenberg, N., Cumberland, A., Spinrad, T. L., Fabes, R. A., Shepard, S. A., Reiser, M., . . . Guthrie, I. K. (2001). The relations of regulation and emotionality to children's externalizing and internalizing problem behavior. *Child Development*, *72*(4), 1112–1134. doi:10.1111/1467-8624.00337

Eisenberg, N., Fabes, R. A., Guthrie, I. K., & Reiser, M. (2000). Dispositional emotionality and regulation: Their role in predicting quality of social functioning. *Journal of Personality and Social Psychology*, *78*(1), 136–157. doi:10.1037/0022-3514.78.1.136

Eisenberg, N., Zhou, Q., Spinrad, T. L., Valiente, C., Fabes, R. A., & Liew, J. (2005). Relations among positive parenting, children's effortful control, and externalizing problems: A three-wave longitudinal study. *Child Development*, *76*(5), 1055–1071. doi:10.1111/j.1467-8624.2005.00897.x

Else-Quest, N. M., Hyde, J. S., Goldsmith, H. H., & Van Hulle, C. A. (2006). Gender differences in temperament: A meta-analysis. *Psychological Bulletin*, *132*(1), 33–72. doi:10.1037/0033-2909.132.1.33

Espy, K. A. (1997). The shape school: Assessing executive function in preschool children. *Developmental Neuropsychology*, *13*(4), 495–499. doi:10.1080/87565649709540690

Friedman, N. P., Miyake, A., Young, S. E., DeFries, J. C., Corley, R. P., & Hewitt, J. K. (2008). Individual differences in executive functions are almost entirely genetic in origin. *Journal of Experimental Psychology: General*, *137*(2), 201. doi:10.1037/0096-3445.137.2.201

Hongwanishkul, D., Happaney, K. R., Lee, W. S., & Zelazo, P. D. (2005). Assessment of hot and cool executive function in young children: Age-related changes and individual differences. *Developmental Neuropsychology*, *28*(2), 617–644. doi:10.1207/s15326942dn2802_4

Hu, L. T., & Bentler, P. M. (1999). Cutoff criteria for fit indexes in covariance structure analysis: Conventional criteria versus new alternatives. *Structural Equation Modeling: A Multidisciplinary Journal*, *6*(1), 1–55. doi:10.1080/10705519909540118

Hughes, C., Ensor, R., Wilson, A., & Graham, A. (2009). Tracking executive function across the transition to school: A latent variable approach. *Developmental Neuropsychology*, *35*(1), 20–36. doi:10.1080/87565640903325691

The jamovi project. (2019). *Jamovi* (Version 1.1) [Computer software]. Retrieved from https://www.jamovi.org

Kochanska, G., & Knaack, A. (2003). Effortful control as a personality characteristic of young children: Antecedents, correlates, and consequences. *Journal of Personality*, *71*(6), 1087–1112. doi:10.1111/1467-6494.7106008

Kochanska, G., Murray, K., Jacques, T. Y., Koenig, A. L., & Vandegeest, K. A. (1996). Inhibitory control in young children and its role in emerging internalization. *Child Development*, *67*(2), 490–507. doi:10.2307/1131828

Kochanska, G., Murray, K. T., & Harlan, E. T. (2000). Effortful control in early childhood: Continuity and change, antecedents, and implications for social development. *Developmental Psychology*, *36*(2), 220–232. doi:10.1037/0012-1649.36.2.220

Lee, K., Ng, S. F., Pe, M. L., Ang, S. Y., Hasshim, M. N. A. M., & Bull, R. (2012). The cognitive underpinnings of emerging mathematical skills: Executive functioning, patterns, numeracy, and arithmetic. *British Journal of Educational Psychology*, *82*(1), 82–99.

Lehto, J. E., Juujärvi, P., Kooistra, L., & Pulkkinen, L. (2003). Dimensions of executive functioning: Evidence from children. *British Journal of Developmental Psychology*, *21*(1), 59–80. doi:10.1348/026151003321164627

Liew, J. (2012). Effortful control, executive functions, and education: Bringing self-regulatory and social-emotional competencies to the table. *Child Development Perspectives*, *6*(2), 105–111. doi:10.1111/j.1750-8606.2011.00196.x

Lin, B., Liew, J., & Perez, M. (2019). Measurement of self-regulation in early childhood: Relations between laboratory and performance-based measures of effortful control and executive functioning. *Early Childhood Research Quarterly*, *47*, 1–8. doi:10.1016/j.ecresq.2018.10.004

Metcalfe, J., & Mischel, W. (1999). A hot/cool-system analysis of delay of gratification: Dynamics of willpower. *Psychological Review*, *106*(1), 3–19. doi:10.1037/0033-295X.106.1.3

Miller, M. R., Giesbrecht, G. F., Müller, U., McInerney, R. J., & Kerns, K. A. (2012). A latent variable approach to determining the structure of executive function in preschool children. *Journal of Cognition and Development*, *13*(3), 395–423. doi:10.1080/15248372.2011.585478

Miyake, A., Friedman, N. P., Emerson, M. J., Witzki, A. H., Howerter, A., & Wager, T. D. (2000). The unity and diversity of executive functions and their contributions to complex "frontal lobe" tasks: A latent variable analysis. *Cognitive Psychology*, *41*(1), 49–100. doi:10.1006/cogp.1999.0734

Murphy, B. C., Eisenberg, N., Fabes, R. A., Shepard, S., & Guthrie, I. K. (1999). Consistency and change in children's emotionality and regulation: A longitudinal study. *Merrill-Palmer Quarterly*, *45*(3), 413–444.

Murray, K. T., & Kochanska, G. (2002). Effortful control: Factor structure and relation to externalizing and internalizing behaviors. *Journal of Abnormal Child Psychology*, *30*(5), 503–514. doi:10.1023/A:1019821031523

Neuenschwander, R., Röthlisberger, M., Cimeli, P., & Roebers, C. M. (2012). How do different aspects of self-regulation predict successful adaptation to school? *Journal of Experimental Child Psychology*, *113*(3), 353–371. doi:10.1016/j.jecp.2012.07.004

Nigg, J. T. (2017). Annual Research Review: On the relations among self-regulation, self-control, executive functioning, effortful control, cognitive control, impulsivity, risk-taking, and inhibition for developmental psychopathology. *Journal of Child Psychology and Psychiatry*, *58*(4), 361–383. doi:10.1111/jcpp.12675

Posner, M. I., Rothbart, M. K., & Rueda, M. R. (2008). Brain mechanisms and learning of high level skills. In A. M. Battro, K. W. Fischer, & P. J. Léna (Eds.), *The educated brain: Essays in neuroeducation* (pp. 151–165). New York, NY: Cambridge University Press.

Prencipe, A., & Zelazo, P. D. (2005). Development of affective decision making for self and other: Evidence for the integration of first-and third-person perspectives. *Psychological Science*, *16*(7), 501–505. doi:10.1111/j.0956-7976.2005.01564.x

Putnam, S. P., & Rothbart, M. K. (2006). Development of short and very short forms of the Children's Behavior Questionnaire. *Journal of Personality Assessment*, *87*(1), 102–112.

Putnick, D. L., & Bornstein, M. H. (2016). Measurement invariance conventions and reporting: The state of the art and future directions for psychological research. *Developmental Review*, *41*, 71–90. doi:10.1016/j.dr.2016.06.004

R Core Team. (2018). *R: A language and environment for statistical computing* [Computer software]. Retrieved from https://cran.r-project.org/

Raffaelli, M., Crockett, L. J., & Shen, Y.-L. (2005). Developmental stability and change in self-regulation from childhood to adolescence. *The Journal of Genetic Psychology*, *166*(1), 54–76. doi:10.3200/GNTP.166.1.54-76

Razza, R. A., & Blair, C. (2009). Associations among false-belief understanding, executive function, and social competence: A longitudinal analysis. *Journal of Applied Developmental Psychology*, *30* (3), 332–343. doi:10.1016/j.appdev.2008.12.020

Rothbart, M. K., & Bates, J. (2006). Temperament. In N. Eisenberg, W. Damon, & R. M. Lerner (Eds.), *Handbook of child psychology: Social, emotional, and personality development* (Vol. 3, pp. 99–166). Hoboken, NJ: John Wiley & Sons Inc.

Rothbart, M. K., Posner, M. I., & Kieras, J. (2006). Temperament, attention, and the development of self-regulation. In K. McCartney & D. Phillips (Eds.), *Blackwell handbook of early childhood development* (pp. 338–357). Malden, MA: Blackwell Publishers Ltd.

Rothbart, M. K., & Rueda, M. R. (2005). The development of effortful control. In U. Mayr, E. Awh, & S. Keele (Eds.), *Developing individuality in the human brain: A tribute to Michael I. Posner* (pp. 167–188). Washington, DC: American Psychological Association.

Rothbart, M. K., Sheese, B. E., & Posner, M. I. (2007). Executive attention and effortful control: Linking temperament, brain networks, and genes. *Child Development Perspectives*, *1*(1), 2–7. doi:10.1111/j.1750-8606.2007.00002.x

Samyn, V., Roeyers, H., Bijttebier, P., Rosseel, Y., & Wiersema, J. R. (2015). Assessing effortful control in typical and atypical development: Are questionnaires and neuropsychological measures interchangeable? A latent-variable analysis. *Research in Developmental Disabilities*, *36*, 587–599. doi:10.1016/j.ridd.2014.10.018

Schoemaker, K., Mulder, H., Deković, M., & Matthys, W. (2013). Executive functions in preschool children with externalizing behavior problems: A meta-analysis. *Journal of Abnormal Child Psychology*, *41*(3), 457–471. doi:10.1007/s10802-012-9684-x

Schroeder, V. M., & Kelley, M. L. (2010). Family environment and parent-child relationships as related to executive functioning in children. *Early Child Development and Care*, *180*(10), 1285–1298. doi:10.1080/03004430902981512

Spinrad, T. L., Eisenberg, N., & Gaertner, B. M. (2007). Measures of effortful regulation for young children. *Infant Mental Health Journal*, *28*(6), 606–626. doi:10.1002/imhj.20156

Sulik, M. J., Huerta, S., Zerr, A. A., Eisenberg, N., Spinrad, T. L., Valiente, C., … Sallquist, J. (2010). The factor structure of effortful control and measurement invariance across ethnicity and sex in a high-risk sample. *Journal of Psychopathology and Behavioral Assessment*, *32*(1), 8–22. doi:10.1007/s10862-009-9164-y

Talwar, V., & Lee, K. (2008). Social and cognitive correlates of children's lying behavior. *Child Development*, *79*(4), 866–881. doi:10.1111/j.1467-8624.2008.01164.x

Wiebe, S. A., Espy, K. A., & Charak, D. (2008). Using confirmatory factor analysis to understand executive control in preschool children: I. Latent structure. *Developmental Psychology*, *44*(2), 575–587.

Wiebe, S. A., Sheffield, T., Nelson, J. M., Clark, C. A., Chevalier, N., & Espy, K. A. (2011). The structure of executive function in 3-year-olds. *Journal of Experimental Child Psychology*, *108*(3), 436–452. doi:10.1016/j.jecp.2010.08.008

Willoughby, M., & Blair, C. B. (2016). Measuring executive function in early childhood: A case for formative measurement. *Psychological Assessment*, *28*(3), 319. doi:10.1037/pas0000152

Willoughby, M., Blair, C. B., Wirth, R., & Greenberg, M. (2010). The measurement of executive function at age 3 years: Psychometric properties and criterion validity of a new battery of tasks. *Psychological Assessment*, *22*(2), 306–317. doi:10.1037/a0018708

Willoughby, M., Kupersmidt, J., Voegler-Lee, M., & Bryant, D. (2011). Contributions of hot and cool self-regulation to preschool disruptive behavior and academic achievement. *Developmental Neuropsychology*, *36*(2), 162–180. doi:10.1080/87565641.2010.549980

Wolfe, C. D., & Bell, M. A. (2004). Working memory and inhibitory control in early childhood: Contributions from physiology, temperament, and language. *Developmental Psychobiology: The Journal of the International Society for Developmental Psychobiology*, *44*(1), 68–83. doi:10.1002/dev.10152

Zelazo, P. D., & Müller, U. (2002). Executive function in typical and atypical development. In U. Goswami (Ed.), *Blackwell handbook of childhood cognitive development* (pp. 445–469). Malden, MA: Blackwell Publishers Ltd.

Zelazo, P. D., Müller, U., Frye, D., & Marcovitch, S. (2003). The development of executive function in early childhood. *Monographs of the Society for Research in Child Development*, *68*(3), 1–151. doi:10.1111/j.0037-976X.2003.00261.x

Zhou, Q., Chen, S. H., & Main, A. (2012). Commonalities and differences in the research on children's effortful control and executive function: A call for an integrated model of self-regulation. *Child Development Perspectives*, *6*(2), 112–121. doi:10.1111/j.1750-8606.2011.00176.x

Zoelch, C., Seitz, K., & Schumann-Hengsteler, R. (2005). From rag (bag) s to riches: Measuring the developing central executive. In W. Schneider, R. Schumann-Hengsteler, & B. Sodian (Eds.), *Young children's cognitive development. Interrelationships among executive functioning, working memory, verbal ability and theory of mind* (pp. 39–69). Mahwah, NJ: Lawrence Erlbaum Associates.